INTRODUCTION[1]

David Papineau

The Epistemology of Science

The philosophy of science can usefully be divided into two broad areas. The epistemology of science deals with the justification of claims to scientific knowledge. The metaphysics of science investigates philosophically puzzling features of the world described by science. In effect, the epistemology of science asks whether scientific theories are true, while the metaphysics of science considers what it would tell us about the world if they were.

The essays in this collection will be concerned with the epistemology of science. They will ask whether we are justified in believing scientific theories, and what attitude we should take to them if we can't. Perhaps the metaphysics of science has as much claim to the title of philosophy of science as the epistemology of science. However, it is a heterogenoeus area that is resistant to anthologizing. Problems in the metaphysics of science tend either to merge into general metaphysics (like the analysis of causation, or probability, or laws of nature) or to fall withn the subject area of specific sciences (like questions about quantum indeterminacy, or about the units of of natural selection). The epistemology of science, by contrast, deals with problems which arise for science in general, rather than specific sciences, yet at the same time are distinguishable from problems arising in other areas of philosophy.

What is more, in the last decade or so a number of philosophers have done much to consolidate and unify work within epistemology of science. In the first half of this century, the dominant tradition in the epistemoloy of science was the logical empiricism of Rudolf Carnap and Carl Hempel, which used the techniques of formal logic and mathematics to analyze the structure of scientific theories and to formulate theories of scientific explanation and confirmation (cf. P.H. Nidditch, ed., The Philosophy of Science, Oxford Readings in Philosophy, 1968). However, in the 1960s this logic-based approach was challenged by the historically orientated work of N.R. Hanson, T.S. Kuhn, and Paul Feyerabend, who appealed to detailed case studies from the history of science to argue that the presuppositions of logical empiricism were fatally flawed (cf. Ian Hacking, ed., Scientific Revolutions, Oxford Readings in Philosophy, 1981).

This led to a schism in studies of scientific theorizing. Some philosophers continued to work within the tradition of Carnap and Hempel, formalizing ideal patterns of scientific reasoning. But as many became convinced that this formal approach bore little relation to the reality of scientific practice, and turned instead to historical analysis for insight into the structure of science. One by-product of this historical turn was that many philosophers of science, and even more historians and sociologists of science, became sceptical about any objective standards of scientific rationality, and came to view theory choices as nothing but expressions of social and institutional pressures.

---

[1] I would like to thank Stathis Psillos for discussing this Introduction with me, and also for his help in selecting the material for this volume.

The essays in this collection, along with much other recent work in the epistemology of science, help to bridge this gap between the logical empircists and the historicists. Contemporary epistemologists of science have learned from Kuhn and others that real science is less rule-bound than logical empiricism supposed, but they do not automatically conclude that it lacks rationality altogether. Or, to put the point the other way round, while contemporary epistemologists of science retain the logical empiricists' concern with isues of scientific objectivity, they are no longer committed to the logical empiricists' overly formal account of what objectivity requires.

Realism and its Antitheses

Much recent debate in the epistemology of science centres on the issue of scientific realism. However, discussions of scientific realism, and in particular of the alternatives to realism, are beset by much terminological confusion. Some initial clarification will be helpful.

Suppose we take realism, for any body of putative knowledge, scientific or not, to involve the conjunction of two theses. (1) An independence thesis. Our judgements answer for their truth to a world which exists independently of our awareness of it. (2) A knowledge thesis. By and large, we can know which of these judgements are true.

Realism, as so defined, is threatened by an obvious internal tension: if the world is independent of our awareness of it, then how can we gain secure knowledge of it? Various resolutions of this tension are possible. Realists will seek some way of simultaneously upholding both the knowledge thesis and the independence thesis. But there are also two traditional alternatives to realism, defined by their rejection of one of these theses. The idealist and verificationist tradition abandons the independence thesis, arguing that the very notion of some further world, beyond the world as we perceive it, is incoherent. Sceptics, by contrast, abandon the knowledge thesis, and accept that we cannot know the truth about the world.

In contemporary epistemology of science, scepticism is the main alternative to realism. This marks a contrast with the epistemology of everyday knowledge. When mainsteam philosophers consider our everyday knowledge of objects like trees and tables, they take the most serious alternative to realism to be some version of idealism or verificationism. Thus, for example, "phenomenalists" argue that it is impossible for a concept to stand for anything except a pattern of sensations, and that therefore it is a confusion to think that human judgements can answer anything beyond the world as it appears in sense perception. Given this, phenomenalists can then argue that knowledge of trees and tables is unproblematic, on the grounds that there is no difficulty knowing about our own sensations.

It is true, of course, that scepticism is the first topic in many mainstream epistemology textbooks. But nobody outside the philosophy classroom seriously questions our knowledge of medium-sized physical objects like trees and tables. In mainstream epistemology, sceptical arguments about trees and tables reflect back our assumptions: since we obviously do know about tables, an argument that such knowledge is impossible challenges us to find the flaw in our reasoning.

By contrast, scepticism in the epistemology of science is by no means just a philosophical exercise. For it is not at all obvious that we do know about the entities postulated by modern scientific theories, such as gravitational waves or neutrinos. After all, we never have any direct sensory evidence for these entities. And the track record of past theories which postulated similar unobservable entities is not good. Considerations like these have persuaded a significant number of contemporary philosophers of science that knowledge of such entities really is unattainable.

Scepticism has not always been the main alternative to a realist epistemology of science. In the past many philosophers of science, from J.S. Mill to the Rudolf Carnap of the Aufbau (1928), opted for a phenomenalist account of the content of scientific claims, arguing that terms like "mass", "charge", and "force" were properly understood as standing for complexes of observable circumstances. Indeed this was arguably the dominant view until well into this century. But the rise of modern micro-physics, with its talk of atoms and fields, has made this view unacceptable. Carnap's attempts to reduce such theoretical talk to observational claims ran into technical obstacles. And in any case the view is highly counterintuitive. It is one thing to maintain that claims about trees are really claims about sensations. But it is hard to take this thesis seriously for claims about unobservable objects like electrons.

In consequence, nearly all contemporary philosophers of science accept that science aims literally to describe an unobservable world of microscopic particles and intangible waves. But a significant number conclude from this that science cannot succeed in this aim: since the world science aims to decribe is beyond the reach of human perception, we have no reason to think that its theories are true.

These issues can be obscured by terminology. Contemporary sceptical opponents of scientific realism call themselves "instrumentalists", or "fictionalists" or "constructive empiricists", and would probably object to the simple epithet "sceptic". Even so, and unlike earlier phenomenalist philosophers of science like Mill or Carnap, these contemporary philosophers of science all accept that scientific theories aim literally to portray an unobservable world, and conclude that for this reason that we would be wrong to believe any scientific theories. If these philosophers differ from simple sceptics, it is only in adding to this rejection of belief the further thought that scientifc theories can nevertheless be useful "instruments" or "fictions" for predictive purposes, and to this extent can be "accepted" as working tools.

Another source of terminological confusion is the term "anti-realism". This term was first coined by Michael Dummett to describe a position in the idealist-verificationist tradition. Dummett's "anti-realism" does not aim to construct the world out of sensations, in the style of simple phenomenalism, but even so it insists, with the idealist-verificationist tradition, that our judgements cannot possibly answer to conditions that are beyond the power of humans to verify. On the other hand, philosophers of science, especially in the United States, have come to use the term "anti-realism" to stand for sceptical views, and in particular for the sceptical attitude towards scientific theories that is the most common contemporary alternative to scientific realism.

Everyone can of course use define their terms as they wish, but there is much scope for misunderstanding here. Note in particular that "anti-realism" in Dummett's sense directly contradicts "anti-realism" in the sense of American philosophers of science. Dummett's "anti-realism", like more traditional idealism and verificationism, wants to uphold our claims to knowledge, by arguing that such claims should not be read as answering to a world beyond our ken. By contrast, American "anti-realism" wants to reject any scientific claims to knowledge about the unobservable world, precisely on the grounds that such claims do answer to a world beyond our ken. The only common feature of these two views is their rejection of the conjunction of (1) and (2) that I earlier used to define realism. However, while Dummett's "anti-realism" upholds the knowledge thesis by rejecting the independence thesis, American "anti-realism" does exactly the opposite.

	SCEPTICISM (Constructive empiricism Fictionalism Instrumentalism "Anti-realism" - US)	REALISM	IDEALISM (Verificationism Phenomealism "Anti-realism" - Dummett)
World:	Independent	Independent	Dependent
Knowledge:	No	Yes	Yes

A number of these points about scientific realism are at issue in the first two papers in this collection. Arthur Fine argues in favour of a position he calls "the natural ontological attitude" ("NOA"). This position consists of what he claims are a set of truisms about science common to both "realism" and "antirealism". Central among these truisms, argues Fine, is the doctrine that well-confirmed scientific theories should be accepted as true. Fine argues that the mistake made by both "realists" and "antirealists" is to add overblown metaphysical theses about the nature of truth and reality to their homely core of agreed assumptions.

Alan Musgrave agrees with Fine that NOA is the right attitude to scientific theorizing. But he objects that NOA is itself a species of "realism", not a core that can be agreed between "realists" and "antirealists". After all, as he understands the term, "antirealists" do not accept well-confirmed scientific theories as true. Musgrave cites Bas van Fraassen and Larry Laudan as two prominent contemporary "antirealists" who are quite explicit in their insistence that it is wrong to believe in the truth of any scientific theory about the underlying structure of the unobservable world.

What is more, argues Musgrave, there is no reason why a "realist", as he takes himself to be, should add any dubious metaphysics to NOA's acceptance of theories as true. It is enough for scientific "realism", in Musgrave's sense, that you should accept that science tells us the truth about the unobservable world. Further metaphysical views about the nature of truth and reality are unnecessary additions to this basis "realist" stance.

In large part this apparent dispute between Fine and Musgrave is simply a result of their addressing different debates. Fine is uninterested in the sceptical option, and is instead thinking of the dispute between realism and anti-realism in Dummett's sense, that is, of how far the world itself can be distinguished from the way it appears to human beings. This is the issue that NOA is neutral on. Fine thinks that both the realist, who insists that our judgements answer for their truth to a world of verification-independent facts, and the Dummett-style anti-realist, who denies this, are making heady metaphysical claims that sober philosophers of science would do well to avoid.

Musgrave is happy to agree, but he thinks Fine's NOA is itself a form of realism. This is because he is primarily concerned with scepticism as the alternative to "realism". From his point of view, Fine's NOA, and the two views it mediates between, may as well all be deemed versions of "realism", since they all maintain, against scepticism, that we ought to accept the best scientific theories about unobservables as true.

In terms of the above table, we can say that Fine is concerned with the dispute between the two right-hand columns. He wishes to maintains that we should embrace only those truisms which are common to both. Musgrave, on the other hand, is concerned with the dispute between the left hand column and the two right-hand ones (even though he agrees with Fine in favouring the common core of the right-hand columns against the left).[2]

One obvious moral of this pair of papers is that readers should be wary of terminology. "Realism" and "antirealism" are just two examples. A number of related terms, including "empiricist", "positivist", "pragmatist", and "instrumentalist", vary in meaning as used by different philosophers. Careful readers of the papers in this collection will note how authors introduce such terms and attend to any other indications of the meanings they attach to them.

---

[2] This diagnosis of the debate is not entirely clear-cut. Fine's begins his paper by rejecting a number of arguments which are normally used to defend belief in scientific theories against scepticism, and to this extent at least he seems to see scepticism as the alternative to scientific realism. But later in the paper he appears to switch tack, when he says explicitly that his "antirealists" accept theories as true (this volume, p. [270]).

The Underdetermination of Theory by Evidence

One central challenge to scientific realism comes from the underdetermination of theory by observational evidence. Suppose that two theories T_1 and T_2 are empirically equivalent, in the sense that they make the same observational predictions. Then no of body of observational evidence will be able to decide conclusively between T_1 and T_2.

Note that the problem with such theories is not just that the choice between T_1 and T_2 is underdetermined by current observational evidence. If current evidence fails to decide between two theories, then the obvious response is to suspend belief for the time being, and to seek out experiments that will decide between them. But with genuinely empirically equivalent theories this option is not available. If all the observational predictions of T_1 and T_2 are identical, then there is no experiment that can conclusively eliminate one in favour of the other.

The thesis of the underdetermination of theory by evidence asserts that we will always be left us with empirically equivalent theories, however much evidence we gather. There are two persuasive arguments for this thesis. One stems from the "Duhem-Quine" thesis, which asserts that any theory can retain its central assumptions in the face of any anomalous evidence, by making adjustments to less central assumptions. Suppose we start with two competing theories T_1 and T_2, and look to future evidence to decide between them, as in the last paragraph. Given the Duhem-Quine thesis, it follows that, even after an indefinite amount of future evidence, we will still have two theories T_1' and T_2', derived from the original pair by the successive revisions occasioned by this evidence, and which this evidence therefore fails to decide between.

A more direct argument for the inevitable existence of empirically equivalent theories starts with given T_1, and then points out that we can always gerrymander a different T_2 which makes exactly the same predictions. The simplest version of this strategy will simply take T_2 to assert all the observational claims made by T_1 but to deny the existence of any of the unobservable mechanisms postulated by T_1. More interesting versions of the argument do not just eliminate the unobservable mechanisms postulated by T_1, but replace them with extra "self-correcting" structures designed to yield exactly the same observational appearances. (For example, if T_1 is some dynamic theory, make T_2 the theory that the universe is accelerating at 1ft/sec^2 in a given direction, and add a universal force acting on all bodies to produce this acceleration. The result will be that T_1 and T_2 predict exactly the same observable relative motions.)

The paper by Lawrence Sklar addresses the issues raised by the underdetermination of theory by evidence. Sklar is initially attracted to the thought that any pair of empirically equivalent theories are really just notational variants of each other, the same theory framed in different words, like Newton's Principia written in Latin and English. But he recognizes that this position lacks plausibility. It is in effect a version of the old phenomenalist view of scientific theories (the "positivist" view for Sklar) according to which claims about apparently unobservable entities like electrons are really claims about observable phenomena: for note that empirically equivalent theories will automatically be notational variants only if is impossible to make meaningful (and hence contradictable) claims about any reality behind the observable appearances.

Sklar's paper is concerned to explore the options left open once we accept the underdetermination of theory by evidence (though in the end he is doubtful that any of these options is acceptable). One option is to embrace scepticism, on the grounds that we ought never to believe any theory, if no empirical evidence can conclusively eliminate its empirically equivalent alternatives. Another is to seek ways of discriminating between empirically equivalent theories, arguing that even when a number of theories make the same observational predictions, one can be more belief-worthy than the others.

The paper by Bas van Frassen embraces the former option. According to van Fraassen's "constructive empiricism" (which is elaborated in greater detail in The Scientific Image, 1980) we ought never to believe in the truth of any theory which goes beyond the observable phenomena. At most we should believe that such a theory is "empirically adequate", that is, that it is correct in what it says about the observable portion of the world. Much of van Fraassen's article is concerned to show that this notion of empirical adequacy, and the associated distinction between the observable and the unobservable, is better explained within the "semantic" account of scientific theories, which identifies theories with sets of models, than within the more traditional "syntactic" account of theories as sets of sentences. (In connection with the observable-unobservable distinction, it is worth observing that a sceptic like van Fraassen asks rather less of the observable-unobservable distinction than a phenomenalist like Mill or Carnap: for, where phenomenalists maintain that we cannot meaningfully talk about unobservables, the scientific sceptic holds only that beings with our limited perceptual abilities ought not to believe any such claims.)

The alternative, non-sceptical response to underdetermination is to argue that we can have grounds for believing one among a set of empirically equivalent theories. At first sight this might seem unpromising. If nothing we observe can rule out alternative theories, then how can we possibly be justified in selecting one for belief? But this last argument implicitly imposes very high standards on justified belief. It assumes that we are only entitled to believe the logical consequences of our observations (since it moves immediately from the existence of alternative theories consistent with our observations to the inadmissibility of belief in any particular theory). However, this an unreasonably strong requirement on justified belief. Nor is it one that most contemporary scientific sceptics would wish to impose. For it would imply (as van Fraassen notes at the end of his article) that we were not even entitled to believe in a theory's empirical adequacy, since this in itself requires us to make an inductive leap beyond the logical consequences of our finite stock of observational data.

So perhaps there is room for a non-sceptical response to underdetermination. The idea, roughly, would be that, among the empirically equivalent theories consistent with our observational data, some are better explanations than the others, in virtue of their greater simplicity, or elegance, or unifying power, and that these virtues are indications that those theories are true. (Compare the way in which we restrict ourselves to "projectible" generalizations when we perform simple enumerative inductions from finite data at the observable level.)

A preference for the "best explanation" certainly seems to be part of scientific practice. As we saw above, there are always empirical equivalents for any given theory, either based on "self-correcting" mechanisms, or on the rejection of unobservable mechanisms altogther. But few practising scientists would regard the availability of such clumsy alternatives as a good reason for disbelieving their normal theories. Given that the explanations of the observable data given by normal theories are far more elegant, scientists are generally happy to embrace the normal theories as true.

Still, even if "inference to the best explanation" is part of intuitive scientific practice, it is not necessarily a good form of inference. Sceptical philosophers of science will argue that scientists go astray whenever they commit themselves to the truth of their best explanations. One obvious objection to inference to the best explanation is considered in Peter Lipton's paper 'Is the Best Good Enough?' Suppose we allow that scientists standardly make accurate <u>comparative</u> judgements when they judge that T_1 is more likely to be true than $T_2, \ldots, T_n$. This alone won't ensure they will get at the truth, because there may yet be even more likely theories among those they haven't yet thought of. It would be silly to infer that the "best" explanation is true, if it is only the "best" among those that have been so far considered. In the seventeenth century Newton's theory was far and away the best account of gravitational motion on offer. But that was only because nobody had yet been able to formulate general relativity.

In response to this argument, Lipton suggests that it is sometimes possible to use comparative scientific judgements to reach the absolute assessments that scientific realism requires. For we can always ensure that our survey of theoretical options is exhaustive, by including as a "catch-all" T_{n+1} the negation of all the theories $T_1, \ldots, T_n$ so far thought of. True, scientists will sometimes judge, when they have done this, that T_{n+1} is the most likely to be true, that is, that the truth most likely lies among the theories they haven't yet thought of. But in other cases they will rank T_{n+1} below the theories they have thought of, and hence conclude that their "best" theory is absolutely likely to be true.

<u>The Pessimistic Meta-Induction from Past Falsity</u>

Lipton's paper thus defends inference to the best (of all possible) explanations against one posible objection. However, the underlying questions about the philosophical status of this kind of inference remain. Note that Lipton starts with the assumption that scientists can at least make sound <u>comparative</u> judgements of truth-likelihood. But sceptics need not accept this. More generally, sceptic can invoke the underdetermination of theory by evidence to question any link between explanatory goodness (simplicity, elegance, unificatory power) and truth. If there are always alternative theories consistent with the evidence, then what guarantee can there possibly be that the most explanatory theory will generally be true?

This sceptical challenge raises a number of delicate philosophical issues about truth, rationality, and the onus of argument, to which I shall return in the next section.

But first I shall consider a rather simpler argument against inferring the truth of the best explanation. This appeals not to the subtleties of underdetermination, but to direct empirical evidence of past theoretical failures. If we survey cases where scientists have embraced their best explanations as true, these explanations have normally turned out false: consider, for example, Ptolemiac astronomy, the caloric theory of heat, the ether theory of electromagnetism, and so on. Given this poor past record of best explanations, ought we not to conclude that inference to the best explanation will in general lead us to falsehoods, rather than truths?

This "pessimistic meta-induction from past falsity" lies behind Larry Laudan's 'A Confutation of Convergent Realism'. Laudan attacks the argument that if some scientific theory is "successful", in the sense of generating confirmed predictions across a variety of contexts, then the best of all possible explanations of that success is the truth of that theory ("it would be a miracle if everything worked as the theory predicted, yet the theory were false"). He surveys a wide range of theories which were arguably successful in this sense, and points out that such success is not normally explained by the truth of the theory involved. For nearly all such theories (he gives a long list from history) have been recognized with hindsight to be false.

A possible realist response (anticipated by Laudan) is to argue only for the <u>approximate</u> truth of succesful theories, rather than their unqualified truth. No doubt predictively successful theories are not always true in every precise detail. But perhaps they can still be assumed to be <u>close</u> to the truth, and in general closer to the truth than their predecessors.

One objection to this ploy is that the notion of "approximate truth" is extremely difficult to articulate clearly. A sustained tradition of research on the notion of "verisimilitude" (see the bibliography at the end of this volume) has made it clear that there can be no interest-independent measure of a theory's distance from the truth. However, even apart from this technical difficulty, there is a more obvious objection to the realist appeal to approximate truth. Namely, that most past predictively successful theories are not even close to the truth, on any intuitively plausible reading of "approximate truth". Laudan's objection to realism is not just that past theories turn out to be wrong on points of detail. Rather, they tend to be radically at variance with the truth, commiting themselves to a range of explanatory entities (like crystal spheres, or caloric fluid, or the ether) that have no counterparts whatsoever in reality.

John Worrall explores a different realist response to Laudan's argument. He concedes that predictively successful past theories standardly contain fundamental errors. But he argues that this does not require a blanket rejection of all scientific claims about the unobservable mechanisms behind the observable predictions. In Worrall's view, the lessons of history bear differentially on different components in scientific theories. More specifically, he argues that history shows that past theories are characteristically wrong about the <u>nature</u> of the unobsevable realm, but not about the <u>structure</u> of its behaviour. In his central example, Worrall argues that nineteenth-century scientists were wrong to think of electomagnetic radiation as embodied in an "ether", but quite right about the mathematical equations governing electromagnetism. Worrall draws the general moral that we should believe in the structure of unobservable reality postulated by successful theories, but avoid committing ourselves to any claims about the nature of that reality.

The general strategy exemplified by Worrall's paper seems to offer the best hope for realism. In the face of past theoretical failures, realists need to show that some parts of the failed theories fared better than others. If they can then identify some principled difference between the good parts and the bad parts, they can recommend belief in the good parts of current theories.

It is a further question, however, whether Worrall's specific way of drawing the distinction does the trick. Some philosophers of science would argue that, since our intellectual access to unobservable entities is always mediated by a structure of theoretical assumptions, rather than by direct insight into their nature, Worrall's restriction of belief to structural claims is in fact no restriction at all. (Cf. Psillos, 1995.) If this is right, then realists need to find some better way of distinguishing the parts of theories which are likely to be discredited from the parts which are worthy of belief.

Truth and Rationality

Let us suppose that realism can succeed in blocking the pessimistic meta-induction in some such way. Realists will not then embrace our current theories *in toto* as the best explanation of predictive success, but only those parts which the evidence of history indicates to be genuinely responsible for such success.

However, realism still needs to deal with the other sceptical challenge mentioned earlier. By inferring the best explanation of predictive success, realists suppose that the best (simplest, most elegant, most unifying) explanation is likely to be true. Sceptics challenge this supposition, pointing out that more than one unobservable explanation will always be consistent with any body of observational evidence. Since we have no independent acess to the unobservable realm, how can we possibly know that the "better" expanation is generally the true one?

There are two possible answers to this challenge, corresponding to the middle and right-hand columns in the table given in the first section. Both are "realist" in Musgrave's sense, in that they both reject scepticism and uphold belief in scientific theories. But in Dummett's (and Fine's) sense, one is "realist" and the other "anti-realist", since they disagree on whether truth involves conformity to an independent reality.

The Dummettian "anti-realist" option is illustrated in the paper by Brian Ellis. If we think of truth as correspondence to an independent reality, argues Ellis, then there is no alternative to van Fraassen's scepticism, for there is no way of establishing that theories with explanatory virtues are generally true. However, this gap between explanatory goodness and truth can be bridged, urges Ellis, if we take the truth to be by definition the view we ought rationally to hold given all possible observable evidence. For, if it is a norm of rationality that we should believe the most explanatory (simple, elegant, unifying) among underdetermined theories, then the best such theory (given total evidence) will by definition be a true theory.

Ellis's position, which he characterizes as "internal realism", stands or falls with his "pragmatic" account of truth as whatever is justifiable by rational norms. Ellis in effect takes rational norms as given _a priori_, and then defines truth in terms of these _a priori_ norms. The resulting position is anti-realist in Dummett's sense. It doesn't make judgements answer for their truth to a realm of sensation, in the style of traditional phenomenalist anti-realism, but it does deny that judgements answer to anything independent of the rational norms of human thought.

Ellis's "pragmatic" account of truth is by no means uncontroversial. Note that it is not an _argument_ for such an account of truth that it would deliver an answer to the sceptical challenge. (Analogously, that we _could_ know about trees, if they _were_ just patterns of perceptions, is not a good reason for thinking that trees _are_ patterns of perceptions.) Many philosophers prefer the alternative view, that truth should be defined as conformity to an independent reality, rather than in terms of the norms of human rationality.

It would take us too far afield to discuss the proper analysis of the concept of truth here. Let us instead consider whether the non-pragmatic view of truth has any answer to the sceptical challenge. Ellis says that, if truth is conformity to an independent reality, then there is no way to link explanatory virtue with truth. But this is too quick. A non-pragmatist account of truth will not supply any _conceptual_ link between truth and explanatory virtue, since it does not define truth in terms of explanatory virtue. But it still leaves open the possibility of an _empirical_ connection between explanatory virtue and truth. If such an empirical link could be defended, then this would offer a different way of defending inferences to the best explanation against the sceptical challenge.

The second contribution to this volume by Larry Laudan bears on this possibility. Laudan is not concerned to defend realism against scepticism. But he shows how principles of scientific theory-choice, such as inference to the best explanation, might be vindicated by empirical generalizations which connect theoretical virtues with the goals of science.

Laudan's starting-point is an older debate about the relevance of the history of science to the philosophy of science. The "historical turn" instigated by Hanson, Kuhn, and Feyerabend in the 1960s led philosophers of science to test methodological proposals against empirical evidence from the history of science. If Newton could be shown to have violated Carnap's methodology, say, then this was taken to count against Carnap's methodology. However, the logic of this kind of argument is obscure. Why should a methodologist's precepts have to fit the empirical facts of Newton's practice? As Laudan observes, even if we accept that Newton is a member of the scientific elite, we needn't therewith accept that every aspect of his actual practice is methodologically exemplary.

Laudan offers a different explanation of the relevance of historical evidence to methodology. Suppose we distinguish the _aims_ from the _means_ of scientific practice. Aims may include finding true theories, or predictively reliable theories, or theories that offer proofs of God's existence. Given some such end, x, the practice of choosing theories that display a certain putative virtue, y, can be viewed as a means to that end. This suggests that methodological principles have the form: "if you want x, you ought to choose theories with y". That is, methodological principles have the form of hypothetical imperatives, specifying a means to an end. As such, they can be assessed on empirical grounds, like all similar means-ends recommendations: does the empirical evidence show that y is in fact an effective means to x? More specifically, in the present context, does the historical record show that theories with feature y are a good route to aim x? From Laudan's point of view, Newton's practice is methodologically relevant, not because of Newton's special status, but simply because it provides cases which might help to show whether choosing y generally leads to x.

Though Laudan leaves the question open in this second article, he does not in fact think that truth (as opposed to predictive reliability, say) is a sensible aim for science. Since he defends the pessimistic meta-induction, he thinks that the historical evidence demonstrates the ineffectiveness of any methodological means by which scientists might attempt to reach the truth (cf. Laudan, 1984, p. 137).

However, if we are able to block the pessimistic induction along the lines outlined in the last section, then perhaps we can press Laudan's means-ends view of scientific rationality into the service of realism. Perhaps relevant historical evidence will enable us to identify certain methodological strategies as effective routes to true theories. A methodology for science of this kind would have affinities with the reliabilist tradition in mainstream epistemology, which holds that a belief qualifies as knowledge if it has been produced by a reliable method, that is, by a method which, as a matter of empirical fact, generally produces true beliefs. This tradition in effect inverts the approach exemplified by Ellis: instead of defining truth in terms of a priori norms, it identifies the rationally correct methodology as that which, as a matter of empirical fact, provides an effective route to the truth.

Richard Boyd has developed such an approach to the philosophy of science in a series of articles over the last two decades. Boyd offers an empirical argument for the reliability of modern scientific methods: namely, that the only good explanation for the predictive success of science is that modern science in general provides an effective route to the (approximate) truth. He points out that the procedures by which scientists devise and test new theories are standardly informed by background assumptions provided by already established theories. Boyd maintains that these procedures for developing new theories could scarcely be expected to yield predictively success, unless these established theories are by and large true. (Cf. Lipton, this volume, p. [103].)

In the article reprinted here, Boyd explains how this argument can withstand the apparently sceptical implications of the pessimistic meta-induction. He concludes the article by addressing the charge that his position is circular. Note that Boyd moves from the empirical evidence that science is predictively successful, via an inference to the best explanation, to the conclusion that modern scientific method is an effective route to the truth. But the legitimacy of this kind of inference is precisely what his sceptical opponents deny. The sceptic's original challenge was to the legitimacy of first-order inferences to the best explanation, such as the inference that the atomic theory of matter is needed in order to explain the data of chemical experiment. Boyd is now in effect responding to this sceptical challenge with a meta-inference of the same form: the best explanation of science's general predictive success is that modern science generally gets at the truth. Boyd's opponents have not been slow to object that this begs the issue. After all, if they don't accept that the atomic theory of matter is needed to explain the chemical data, why should they accept science's truth is needed to explain its predictive success? (Cf. Fine, this volume, [p. 262-3]; Laudan, 'A Confutation of Convergent Realism', this volume, [p. 240-1].)

Boyd's response is that this last inference is not to be considered in isolation, but as part of an overall "realist package". This package should be compared as a whole with the overall alternatives. If the realist package is superior as a whole to the alternatives, then the charge of cicularity falls away. Realists of Boyd's stripe take themselves to have an answer to the sceptical challenge to the link between explanatory virtue and truth. They say the predictive success of science is good empirical evidence for this link. True, it is only good evidence by realist lights. But if realism is the right overall package, then these are the right lights.

Philosophers like Laudan and Boyd advocate a naturalized philosophy of science. Instead of seeking to identify principles of scientific rationality on some a priori basis, they look to empirical information about the effectiveness of different scientific practices to decide the right methodology for science. It is worth observing that one important by-product of this "naturalist turn" is the prospect of a reconciliation between philosophers and sociologists of science.

The last two decades have been marked by an explosion of exciting work in the sociology of science. Traditionally, sociologists were content to study the external aspects of science, like the growth of scientific institutions or the structure of scientific education. But recent sociological work has turned inwards to scientific theorizing itself, and aims to show how social influences and interactions play a decisive role in the resolution of specific theoretical debates.

This kind of sociology has been widely regarded as undermining any epistemological analysis of science, on the grounds that epistemology deals in a priori standards of rationality, while sociological studies seem to show that scientific theory-choices aren't governed in this way at all, but rather by scientific power struggles and opportunistic manoeuvres.

This rejection of epistemology, however, presupposes that epistemology traffics solely in a priori principles of theory evaluation. By contrast, if epistemology of science is conducted in the naturalist mode, then the conflict disappears. Naturalized philosophers of science have no axe to grind for a priori methodological principles over social processes. They can happily accept that theory-choices are often determined by social processes. The only normative question they will then want to ask is whether or not those processes are an effective means to scientific aims. Nor is this a question they can answer by themselves. For it is an empirical question, not an a priori one, and the philosophers will therefore need the help of the sociologists and historians of science in to answer it. (For further reading on this issue, see the sections on naturalized philosophy of science and sociology of science in the selected bibliography at the end of this volume.)

Confirmation Theory and Bayesianism

It is becoming increasingly common for issues in the epistemology of science to be discussed within the framework of Bayesian confirmation theory. Confirmation theory seeks to quantify questions of theoretical belief. Instead of simply asking whether we ought to believe theory T given evidence E, confirmation theorists ask how much we ought to believe T given E.

Bayesian confirmation theorists argue that such questions are best formulated as questions about subjective probabilities. Suppose that you attach a initial subjective probability Prob(H) to hypothesis H, and a conditional probability of Prob(H/E) to H on the supposition that evidence E is true. Then E will confirm H for you if your Prob(H/E) is greater than your initial Prob(H). Accordingly, when you observe E you should increase your probability for H to your old Prob(H/E).

A simple theorem of the probability calculus ("Bayes' theorem") states that Prob(H/E)=Prob(E/H)/Prob(E) x Prob(H). This means that E confirms H for Bayesians -- that is, Prob(H/E) is greater than Prob(H) -- if Prob(E/H) is greater than Prob(E).

This makes intuitive sense. If datum E is surprising in itself (Prob(E) is low), but is what you would expect if hypothesis H were true (Prob(E/H) is high), then E is surely good support for H. The apparent bending of light near the sun is a striking fact in itself, but is just what is predicted by the theory of general relativity. So its observation in 1919 gave strong support to Einstein's theory.

Not all philosophers who work on confirmation theory are Bayesians. Clark Glymour, for one, has developed an alternative "bootstrapping" account of confirmation in his Theory and Evidence (1980). In the extract reprinted in this volume, however, he concentrates on criticisms of the Bayesian alternative. Glymour first explains the basic structure of the Bayesian approach (and in particular the "Dutch book argument" for supposing that degrees of belief are a species of probability) and then points to various weaknesses.

For Glymour, the central flaw in Bayesinism is the subjectivity of its probabilities. Bayesianism places no substantial constraints on the probabilities that enter into confirmation relations, other than that they reflect actual degrees of belief. How likely is H, or E, or E given H? For the Bayesian these are questions about the subjective attitudes of individuals.

Glymour argues that this gives Bayesianism a spurious plausibility. Since Bayesianism per se does not restrict the probabilities that enter into confirmation relations, it can always postulate as "reasonable" any probabilities it needs to replicate standard forms of scientific reasoning. Glymour objects that this does not explain the worth of those forms of reasoning (after all, just the same trick could replicate any other forms of reasoning). He also criticizes specific attempts made by Bayesians to identify additional a priori principles governing reasonable degrees of belief.

Wesley Salmon's 'Rationality and Objectivity in Science or Tom Kuhn meets Tom Bayes' offers a naturalistic response to the objection that Bayesian probabilities are arbitrary. Instead of seeking a priori principles governing reasonable subjective probabilities, Salmon suggests we should look to the history of science to tell us about the empirical frequencies with which hypotheses of various kinds have turned out to be successful.

Salmon argues that this move can effect a reconciliation between the formalist tradition of Carnap and Hempel and the historicist approach instigated by Kuhn's The Structure of Scientific Revolutions (1962). Kuhn argues that scientists' judgements of plausibility play a crucial role in deciding their attitude to hypotheses. Many of Kuhn's followers (but not Kuhn himself) have inferred that scientific judgements are consequently arbitrary. Salmon suggests that this charge or arbitrariness is best rebutted by construing such judgements of plausibility as Bayesian probabilities, and then grounding these probabilities in the empirical frequencies with which the hypotheses of the kind in question have proved successful.

It is worth observing that Salmon is open to a charge of circularity analogous to that brought against Boyd. Just as Boyd appealed to empirical evidence concerning scientific practice to vindicate scientists' preference for certain kinds of explanation, so Salmon appeals to similar evidence to vindicate the use of Bayesian reasoning. And, just as Boyd's vindication was a special case of the kind of inference he was trying to defend, so can Salmon's inference, from the empirical evidence to a conclusion about probabilities, be viewed as a special case of Bayesian reasoning. It would be an interesting exercise to consider whether Boyd's reply to this charge of circularity will work for an objectivist Bayesian like Salmon.

Are Theories Important?

Most of the articles in this collection are concerned with the truth or otherwise of scientific theories. However, some contemporary philosphers of science have argued this is the not necessarily the crux of debates about scientific realism. Nancy Cartwright, in How the Laws of Physics Lie (1983), and Ian Hacking, in Representing and Intervening (1983), have elaborated a position that they call "entity realism" as opposed to "theory realism".

This view recognizes that the modern physical sciences are responsible for any number of striking physical effects, from lasers and optic fibres to electron microscopres and superconductors. It denies, however, that these physical effects provide convincing support for fundamental physical theory. Cartwright argues that the standard "derivations" of these effects from fundametal theory are mediated by ad hoc assumptions, mathematical short cuts, and fudge factors. Since these derivational devices are in general unmotivated by the fundamental theory, but are essential to any mathematical account of the relevant physical effects, they fail to provide any inductive support for the fundamental theory. The basic theory effectively does no work in the derivations, by comparison with the simplifications, argues Cartwright, and so deserves no credit.

However, this does not imply the non-existence of the sub-atomic particles and other unobservable entities implicated in the relevant physical effects. For Cartwright and Hacking, the success of our attempts to manipulate these entities is ample testimony to their existence. If these entities were not real, we could not use them to produce physical effects. Hacking encapsulates the view in his well-known slogan: "If you can spray them, they exist."

Entity realism presents an interesting alternative to orthodox theory-based based realism. But it faces some obvious challenges. For a start, it is possible to doubt Cartwright's claim that fundamental physical theory is effectively redundant in the mathematical analysis of physical effects. Certainly simplifications play an important role in such analyses. But many of these simplifications are themselves guided by theoretical considerations, and so arguably add to the credit of the theory when they succeed. Another challenge is to the notion that we can separate commitment to entities from commitment to theories. To a large extent we think about unobservable entities in science as entities that play certain theoretical roles. This makes it difficult to see how we can accept the entities without accepting at least some of the surrounding theory.

In the paper reprinted in this volume, Cartwright takes a somewhat less radical tack. Here she is prepared to concede that the abstract laws of fundamental physics receive inductive support from physical effects. However, she denies that such laws need be universally applicable. Even if they accurately characterize the behaviour of physical phenomena in certain precisely controlled laboratory contexts, it does not follow that they govern all phenomena.

Everyday systems may dance to their own tunes, independently of the forces and equations of basic physics. Cartwright observes that the possibility of such "emergent" patterns is a familiar theme in biological thinking. But she wants to go beyond this, and deny that even physical systems need be governed by fundamental physical laws. Perhaps the behaviour of a falling dollar bill escapes the laws of physics as much as the behaviour of a biological organism. Cartwright urges that we replace the reductionist picture of a unified system founded on a few basic laws with a patchwork of many laws each of limited range.

One issue raised by Cartwright's paper is whether different philosophical morals may apply in different areas of science. Perhaps we should be fundamentalists for physics but not for biology. Or perhaps we should be theory realists in chemistry, entity realists in geology, and outright sceptics in paleobiology. At the beginning of this Introduction I said that the epistemology of science deals with problems which arise for science in general. Certainly most of the pieces in this collection have sought morals that will apply across the scientific board. But perhaps a more fine-grained approach would be worth the extra effort. Now we are clear about the epistemological options on offer, there is no obvious reason why we should expect the same alternative to apply to every scientific discipline.

References

Carnap, R. (1928). Der logische Aufbau der Welt. English version (1967). The Logical Structure of the World. Berkeley: University of California Press.

Cartwright, N. (1983). How the Laws of Physics Lie. Oxford: Clarendon Press.

Glymour, C. (1980). Theory and Evidence. Princeton: Princeton University Press.

Hacking, I. (ed.) (1981). Scientific Revolutions. Oxford: Oxford University Press.

Hacking, I. (1983). Representing and Intervening. Cambridge: Cambridge University Press.

Kuhn, T. (1962). The Structure of Scientific Revolutions. Chicago: University of Chicago Press.

Laudan, L. (1984). Science and Values. Berkeley: University of California Press.

Nidditch, P. (ed.) (1968). Philosophy of Science. Oxford: Oxford University Press.

Psillos, S. (1995). 'Is Structural Realism the Best of Both Worlds?', Dialectica, 49.

Van Fraassen, B. (1980). The Scientific Image. Oxford: Clarendon Press.